Optimization Deployment of Roadside Units with Mobile Vehicle Data Analytics

1st Xuemei Cao
Beijing University of Posts and Telecommunications
Beijing, China
caoxm@bupt.edu.cn

2nd Qimei Cui
Beijing University of Posts and Telecommunications
Beijing, China
cuiqimei@bupt.edu.cn

3rd Sihai Zhang
University of Science and Technology of China
Hefei, China
shzhang@ustc.edu.cn

4th Xueying Jiang
Beijing University of Posts and Telecommunications
Beijing, China
jiangxy@bupt.edu.cn

5th Ning Wang
Beijing University of Posts and Telecommunications
Beijing, China
wangning@bupt.edu.cn

***Abstract*—Mobile self-organizing networks, such as vehicular ad-hoc network (VANET), in most cases rely on infrastructure deployment to provide access to internet services and other resource. Therefore, it is crucial to optimize the deployment of roadside units (RSUs) in vehicle network. In this paper, we propose a RSU optimized deployment scheme based on large vehicle data, which considers the deployment cost and latency performance synthetically. We deploy the RSU problem as a multiobjective optimization problem for mathematical modeling. On this basis, two-step solution is proposed: firstly, the RSU candidate positions can be obtained by considering the road topology and large vehicle data; secondly, the branch and bound algorithm is used to obtain the better RSU deployment based on the mathematical model. The simulation results show that the proposed RSU deployment scheme uses a small amount of RSU can achieve high coverage and greatly reduce the deployment cost. Moreover, the low latency performance of the vehicle access network and the quality of the latency sensitive application service can also be ensured.**

***Index Terms*—RSU, big vehicle data, deployment, cost, latency**

I. Introduction

A. Motivation

The advent of autonomous vehicles and arrival of large data age make the network more intelligent and gradually develop into an intelligent transportation system (ITS) [1]. ITS is less tolerant of latency since information such as vehicle collision should be transmitted as soon as possible. As a result, more and more attention is paid on reducing latency and improving user experience which became a research hotspot in internet of vehicles. The urban vehicular ad hoc networks (VANETs) are considered as an important part of the next generation ITS to alleviate serious traffic problems. To reduce latency and improve quality of service, the infrastructure of VANETs should be optimally deployed. Thus sufficient road-side units (RSUs) deployment is required.

B. Related Work and Contributions

Developments in transport and communications technology made it possible to exchange information between moving vehicles on the roads [2]. Through communications, VANETs was created which consist of a group of highly mobile nodes. VANETs has three possible modes of communication: vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) or both, called V2X [3]. The authors concentrated on spatial point process modeling for random vehicle locations To analyze the performance of V2V communication for VANETs [4] [5]. The communication capability of V2I which is related to the deployment of infrastructure, has great influences on network reliability and information exchange. In this paper, we focus on the deployment of RSUs which acts as wireless access points in VANETs.

The RSU is a fixed node positioned near or along the road to connect with the on-board ad-hoc network [6]. The transmit power of RSU is higher than the mobile station because of wider coverage. Therefore, the RSU deployment can be used to create a network based on vehicles to V2I communications which is feasible for non-interactive applications. In recent years, several different approaches for the deployment of the RSUs have been proposed. In [7], the authors tried to improve data collection and delivery rates in road scenarios by studying the size of the gap between RSU. They initially distributed RSUs along the road and other distances. While in the freeway scene, it was recommended to use a uniform initial RSU distribution (each twice radio range distance) to improve the efficiency of collecting data [8]. However, the uniform distribution was not an optimal initial distribution because it did not take into account the density of vehicles on the road and will be likely to increase the transmission latency. [9] modeled RSU deployment problem as the maximum coverage with time threshold problems (MCTTP [10]) from the perspective of reducing transmission latency. They used genetic heuristic algorithms to solve MCTTP and obtained the best RSU layout. In paper [11], the mobile machines connect to multiple APs and HPN by coordinated multi-point transmission (CoMP) concurrently to achieve ultra-reliable and low-latency communication. In addition, [12] considered the junction position as a potential location for deploying RSU to reduce the latency in the transmission of security messages. More avaliable points were considered in [13]. [14] proposed a cost-efficient RSUs deployment scheme to guarantee that On-Board Units (OBUs) at any place could communicate with RSUs in certain driving time. In addition to considering transmission latency, [15] also took the cost of deployment into account. In [16], the authors found the best positions to place RSUs although they do not optimize the number of RSUs. [17] investigated the RSU deployment problem through transforming the road map into a new graph. However, their method required the road to present horizontal or vertical direction which was not consistent with the actual road topology.

In this paper, a new RSU deployment strategy is proposed which not only improves the quality of time-sensitive services but also reduces deployment cost. The problem of optimizing RSU deployment is modeled mathematically as a multiobjective function. We propose a two-step solution for the optimization objective. At first, the initial RSU deployment location is obtained by considering road topology and analyzing big vehicle data. Secondlly, better deployment locations are confirmed by the branch and bound algorithm. The rest of this paper is organized as follows: Section II gives the mathematical modeling of the RSUs optimal deployments. Section III describes the main steps of the resolution and the performances of our solution are evaluated in Section IV. Finally, section V summarizes and concludes this paper.

II. Big Traffic Data and optimization problem modeling

In this part, we introduce the large-scale vehicle trajectory data which is used in this paper, and give the mathematical model of RSU optimal deployment considering deployment cost and latency performance simultaneously.

A. Data Sets

In collecting Beijing trace, we used the mobility tracklogs obtained from 12509 Beijing taxis carrying GPS receivers during November 2012. Taxi data are chosen because taxis are more sensitive to urban environments in road topologies, and they have a wider range of space and operating time than buses and private cars. The dataset in Beijing is composed of all the location of 12509 taxis for 27 days (from $2012-11-01$ to $2012-11-27$). The dataset contains 785.4 million pieces of information, each one representing a taxi location and being composed of the following attributes:

- Taxi ID: unique identification of each car;
- Timestamps: the format is yyyymmddhhnnss. And y, m, d, h, n, s represent year, month, day, hour, minute, second respectively;
- The location of taxis: a GPS position (latitude and longitude) of the taxi.

The data format for vehicles are shown in TABLE I.

TABLE I: The Format of Beijing Data Set

Taxi ID	Timestamps	Longitude	Latitude
155940	20121101001442	116.3032990	39.941256

B. Optimization problem modeling

The optimized locations for placing RSU are considered based on both the latency performance and the deployment cost. The road system is modeled as a weighted undirected graph $G=(V,E)$, where V is the intersections set of the road system and E is the segment of the road between two intersections. The weight of the edge $e\in E$ is the length of the road segment. Let $X=\{x_1,x_2,...,x_l\}$ be the set of candidate positions to deploy RSUs. A reasonable RSU deployment scheme should meet two conditions: (i) minimizing the transmission delay between the vehicle and the RSU; (ii) minimizing the installation cost of the RSU as much as possible. These two requirements can be expressed by the equation (1).

$$F=\alpha F_1+\beta F_2 \tag{1}$$

where α and β represent the weight of each sub-function. F_1 represents the RSUs installation cost and F_2 represents the delay to reach the closest RSU. We choose $\alpha=1/3$, $\beta=2/3$ as the delay constraint is the most important.

$$F_1=\min\sum_{i=1}^{l}a_ix_i \tag{2}$$

$$F_2=\min\frac{1}{N}\sum_{i=1}^{l}\sum_{k=1}^{N}t_{ki}x_i \tag{3}$$

where t_{ki} is the transmission latency it takes for the vehicle k to reach the nearst coverage of RSU i. It can be calculated as:

$$t_{ki}=\frac{S_n}{W\log(1+\frac{P_nd_{ki}^{-\gamma}}{\sigma^2})} \tag{4}$$

Where W is channel bandwith, σ^2 is the Gaussian noise, S_n is the size of message, P_n is the transmit power, γ

is the path loss and d_{ki} denotes distance which vehicle k reach the nearest RSU i. According to equation (2) and (3), the multiple objective optimization problem is formulated as the following:

$$\begin{aligned} F &= \min \alpha \sum_{i=1}^{l} a_i x_i + \beta \frac{1}{N} \sum_{i=1}^{l} \sum_{k=1}^{N} t_{ki} x_i \\ s.t. \quad & \sum_{i=1}^{l} x_i \leq l \end{aligned} \tag{5}$$

where $x_i \in \{0,1\}$ denotes the optimization variable for the selection of RSU i, a_i denotes the cost of deploying RSU i, l is the number of potential locations for RSU deployment, N is the number of vehicles.

III. Obtain optimized deployment locations for RSU

In our work, the objective is to find the optimal number of RSU to deploy and the optimal RSU location while meeting application latency requirements and deployment cost constraints. The optimization objective function is described in section II, we take two steps to solve this problem. The fist step is to select potential locations in the map to deploy RSU. And the optimal RSU deployments are obtained in the second step according to the branch and bound algorithm.

A. First step: Initialize RSU candidate locations

At this stage, we calculate the initial candidate positions for RSU. In contrast with related work that chooses initial positions randomly, uniformly in intersections, we study the area of road topology and large vehicle data to find candidate RSU locations.

1) Considering Candidate RSU Locations from Road Topology: The road system is modeled as a weighted undirected graph $G = (V, E)$, where V is the set of the intersections of the road system. The actual intersections information $V = \{v_1, v_2, .., v_{n_v}\}$ can be obtained from the OpenStreetMap. Fig.1 shows the actual road topology, where many intersections associate with other roads in a short distance. However, it is not reasonable to set all intersections as alternative locations. Therefore, K-nearest neighbor algorithm [18] is used to deal with V and the Euclidean distance is applied to compute the similarity of the samples.

$$D = \sqrt{(v_i - v_j)^2} \tag{6}$$

It is found that different road grades have different number of intersections in fact at the same intersection. Therefore, according to the road grade to choose a different K values, and choose one from the final K adjacent intersections as a candidate location.

It is an obvious fact that if a junction connects more roads, it is more likely to be used by vehicles. This aspect refers to another type of centrality known as degree centrality. In graph theory, this kind of centrality reflects the capacity of a node to intercept information owing from neighbors. For a graph $G = (V, E)$ with n vertices, the degree centrality of a node v [19] is dened as:

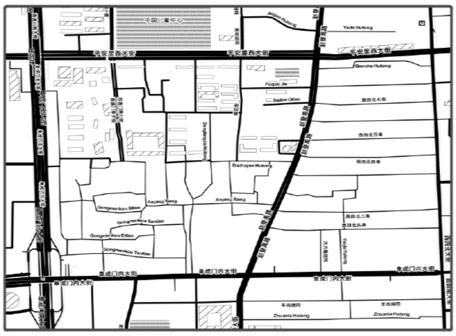

Fig. 1: The real road topological diagram of an intersection.

$$C_D(v) = deg(v)/(l-1) \tag{7}$$

where $deg(v)$ represents the number of connections of the considered node, $deg(v_i) = \sum_{j=1}^{l} g_{ij}(i \neq j)$, $g_{ij} \in \{0,1\}$, if the node v_i to node v_j has direct contact then $g_{ij} = 1$, otherwise $g_{ij} = 0$. In vehicular environment, its importance stems from the obvious fact that the more a junction has incident routes, the more cars are intended to go through it. Therefore, the crossroads with higher degree centrality as best candidate locations.

2) Considering candidate RSU locations based on large vehicle data: This is reasonable and necessary, with the deployment of RSU at vehicle-intensive road crossings. Fig.2 is an average hour vehicle density map for an area (the longitude from 116.3483 to 116.3677 and the latitude from 39.9203 to 39.9330) in Beijing. The heatmap in Fig.2 reveals that some intersections do not have traffic flow, so these intersections will not be candidate RSU locations.

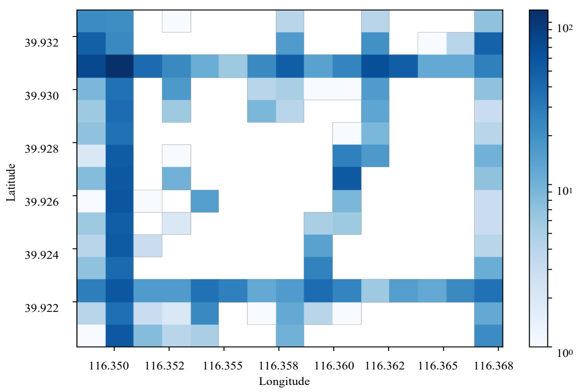

Fig. 2: Vehicle density heat map for an area (the longitude from 116.3483 to 116.3677 and the latitude from 39.9203 to 39.9330).

According to the description, $\lambda_{ki} \in \{0,1\}$ denotes whether vehicle k is in the coverage of RSU i (i.e., $\lambda_{ki} = 1$

when vehicle k is in the coverage of RSU i, otherwise $\lambda_{ki} = 0$). Then we can get a matrix as follows:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1l} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2l} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{N1} & \lambda_{N2} & \cdots & \lambda_{Nl} \end{bmatrix} \tag{8}$$

Each column of this matrix represents the sum of vehicles under a candidate RSU. Therefore, a set point i which satisfies $\sum_{k=1}^{N} \lambda_{ki} \geq 1$ consider as the candidate deployment location for RSU.

Algorithm 1 Branch and bound for optimal RSU deployment objection

Input: Objective function $F(\boldsymbol{x})$ and constraint conditions.
Output: $Output$

```
Initialization optimal solution: F = ∞
while Fathomed nodes = True do
    Selectes a node from fathomed nodes
    Calculates the lower bound value(LB) of each new branched node
    if Feasible solution of minimum lower bound value F* is less than F then
        F = F*
        x = x*
    end if
end while
Output = F, The corresponding integer solution is x.
```

B. Second step: Optimal RSU Placements

As mentioned above, our solution takes into consideration both the application latency requirement and the deployment cost related to RSU deployment. The number of potential locations for deploying RSU is l. There are 2^l possible RSU deployments, since each of the potential locations can be deployed or not deployed with an RSU. So naturally it comes down to a integer planning problem. Branch and bound algorithm [20] which is more effective than enumeration method, is utilized by us to solve this integer programing problem, and the detail is introduced in Algorithm 1.

IV. Results and Discussion

In order to further validate the deployment method, this paper sets up a specific application scenario, namely the above selected areas(the longitude from 116.3483 to 116.3677 and the latitude from 39.9203 to 39.9330).

A. Candidate RSU Locations

We use OSMNx to obtain all intersections information from the OpenStreetMap according to the selected area. Road as shown in Fig.3. The orange points are the intersections of the road, the black connecting lines are roads. Fig.3 shows the fact that there are usually two intersections that are close to each other because many roads in Beijing are parallel. According to the previous analysis, the K-nearest neighbor method is adopted to handle the intersections. Then we consider placing RSU where it should cover the roads as much as possible, so remove junctions where the degree centrality of a crossroad is less than $2/(l-1)$. Of course, intersections (such as certain alleys in residential areas) that have not been crossed by vehicles for a long period of time are also out of the range of candidate locations. The point i which satisfies $\sum_{k=1}^{N} \lambda_{ki} \geq 1$ consider as the candidate deployment location for RSU. After the above steps, the initial RSU deployment is available, as shown in Fig.4.

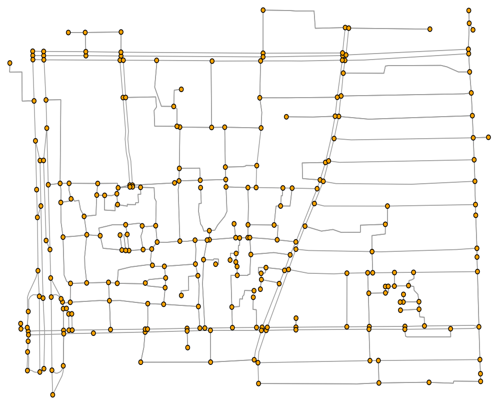

Fig. 3: Road topology map for the selected region. The circles are intersections and solid lines are roads.

Fig. 4: Node diagram for candidate RSU locations. The circles are candidate locations of RSUs and solid lines are roads.

B. Performance of optimized RSU deployment

The deployment costs at all location are set as 1. Fig.5 shows the optimal RSU deployment options with RSU radius of 100m, where the blue dots are candidate RSU positions, the red dots are optimized after the final selection of the RSU deployment location. It is clear that the optimal deployment location for the RSU is not only the main roads, but also the vehicle-dense entry points. The scene area of the simulation selection is $1.6563km \times 1.4183km$. The best way to deploy regardless of cost is a fully deployed RSU in the region that is seamless and does not overlap. The total number of RSUs required for this deployment is $\frac{1656.3\times1418.3}{\pi r^2}$, where r is the radius length of the RSU. When the radius is 100 meters, 200 meters, 300 meters, the number of RSU required is $\frac{1656.3\times1418.3}{\pi 100^2} \simeq 74$, $\frac{1656.3\times1418.3}{\pi 200^2} \simeq 19$, $\frac{1656.3\times1418.3}{\pi 200^2} \simeq 9$ respectively. In this paper, in the optimization algorithm based on vehicle density, road network analysis and low delay requirements, only 43 RSUs with 100 meters radius are used to complete vehicle full coverage, which greatly saves deployment resources. The number of RSUs that need to be deployed between the two deployment scenarios is compared in Fig.6. We can find that with the decrease of RSU coverage radius, the more obvious the advantage of our algorithm is, the more it means to save cost.

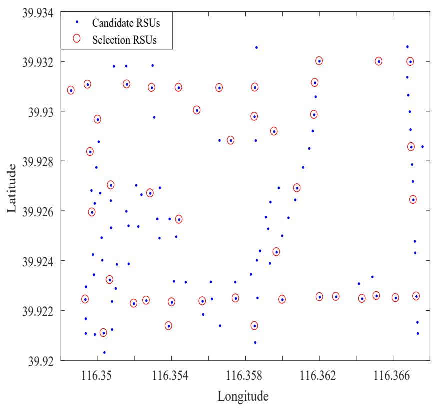

Fig. 5: The alternative deployment location for the RSU and the location actually selected after optimization. The dots are an alternative deployment location for the RSU, and the circles are the one actually selected after optimization.

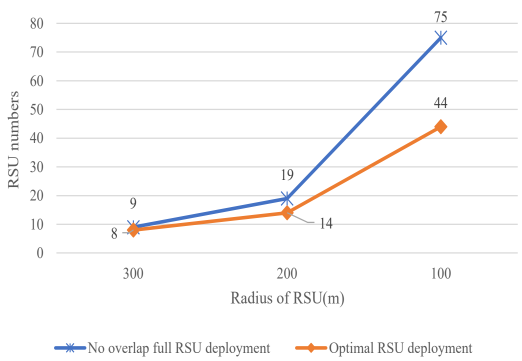

Fig. 6: Contrasting the number of RSUs that need to be deployed.

Fig.7 compares the average latency performance of two different RSU deployment scenarios. And the latency is calculated according to formula (4). Fig.7 shows that our proposed RSU deployment scheme based on large vehicle data considering latency performance and deployment cost can achieve low latency performance with few RSU quantity. So the RSU deployment scheme ensures the low latency performance of the vehicle access network and the quality of the time-sensitive application service with fewer numbers of RSU.

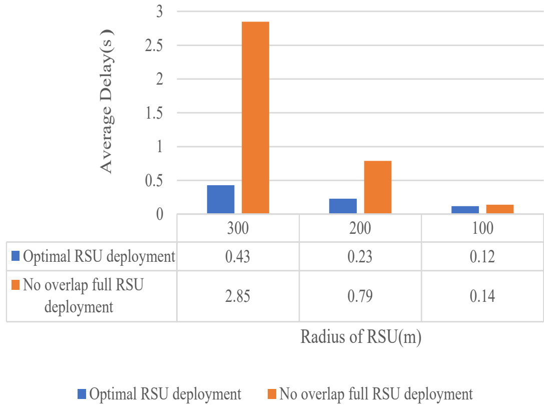

Fig. 7: Comparison of delay between the two deployment scenarios.

V. Conclusion

In this paper, we propose an RSU optimized deployment scheme based on large vehicle data considering deployment cost and latency performance. Firstly, we will use K-nearest neighbor algorithm to remove the overlapping intersection of bidirectional lanes according to the actual road topology situation, and then obtain the position of candidate RSU according to the center degree of intersection nodes and vehicle density selection. According to the requirement of low latency and minimizing deployment cost, the comprehensive objective function is modeled. Branch and bound algorithm is used to solve the problem, and the optimal RSU deployment is obtained. Simulation results show that the proposed optimal deployment scheme uses a small number of RSU to achieve high coverage, thus ensuring the low latency performance of the vehicle access network and the quality of the time-sensitive application service.

Acknowledgement

The work was supported in part by the National Nature Science Foundation of China Project under Grant 61471058 and 61631005, in part by the Hong Kong, Macao and Taiwan Science and Technology Cooperation Projects under Grant 2016YFE0122900, in part by the Beijing Science and Technology Commission Foundation under Grant 201702005 and in part by the 111 Project of China under Grant B16006.

References

[1] Q. Cui, Y. Wang, K. Chen, W. Ni, I. Lin, X. Tao, and P. Zhang, "Big data analytics and network calculus enabling intelligent management of autonomous vehicles in a smart city," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

[2] L. G. Giordano and L. Reggiani, *Vehicular Technologies: Deployment and Applications*. InTech, 2014.
[3] M. Almeida, “Advances in vehicular networking technologies,” 2011.
[4] Q. Cui, N. Wang, and M. Haenggi, “Spatial point process modeling of vehicles in large and small cities,” in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–7.
[5] ——, “Vehicle distributions in large and small cities: Spatial models and applications,” *IEEE Transactions on Vehicular Technology*, 2018.
[6] R. Daher, *Roadside Networks for Vehicular Communications: Architectures, Applications, and Test Fields: Architectures, Applications, and Test Fields*. IGI Global, 2012.
[7] M. Rashidi, I. Batros, T. K. Madsen, M. T. Riaz, and T. Paulin, “Placement of road side units for floating car data collection in highway scenario,” in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2012 4th International Congress on*. IEEE, 2012, pp. 114–118.
[8] B. Aslam and C. C. Zou, “Optimal roadside units placement along highways,” in *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*. IEEE, 2011, pp. 814–815.
[9] E. S. Cavalcante, A. L. Aquino, G. L. Pappa, and A. A. Loureiro, “Roadside unit deployment for information dissemination in a vanet: An evolutionary approach,” in *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation*. ACM, 2012, pp. 27–34.
[10] O. Trullols, J. M. Barcelo-Ordinas, M. Fiore, C. Casetti, and C.-F. Chiasserini, “A max coverage formulation for information dissemination in vehicular networks,” in *Wireless and Mobile Computing, Networking and Communications, 2009. WIMOB 2009. IEEE International Conference on*. IEEE, 2009, pp. 154–160.
[11] S. C. Hung, H. Hsu, S. M. Cheng, Q. Cui, and K. C. Chen, “Delay guaranteed network association for mobile machines in heterogeneous cloud radio access network,” *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2018.
[12] F. Amjad, “Optimal roadside units placement in urban areas for vehicular networks,” in *IEEE Symposium on Computers and Communications*, 2012, pp. 423–429.
[13] Y. Liang, H. Liu, and D. Rajan, “Optimal placement and configuration of roadside units in vehicular networks,” in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. IEEE, 2012, pp. 1–6.
[14] Y. Sun, X. Lin, R. Lu, X. Shen, and J. Su, “Roadside units deployment for efficient short-time certificate updating in vanets,” in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.
[15] S. Mehar, S. M. Senouci, A. Kies, and M. M. Zoulikha, “An optimized roadside units (rsu) placement for delay-sensitive applications in vehicular networks,” in *Consumer Communications and NETWORKING Conference*, 2015, pp. 121–127.
[16] W. H. Ho, K. K. Leung, and J. W. Polak, “Stochastic traffic and connectivity dynamics for vehicular ad-hoc networks in signalized road systems,” in *IEEE Conference on Local Computer Networks*, 2013, pp. 332–335.
[17] P.-C. Lin, “Optimal roadside unit deployment in vehicle-to-infrastructure communications,” in *ITS Telecommunications (ITST), 2012 12th International Conference on*. IEEE, 2012, pp. 796–800.
[18] J. A. Hartigan and M. A. Wong, “Algorithm as 136: A k-means clustering algorithm,” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100–108, 1979.
[19] A. Kchiche and F. Kamoun, “Access-points deployment for vehicular networks based on group centrality,” in *New Technologies, Mobility and Security (NTMS), 2009 3rd International Conference on*. IEEE, 2009, pp. 1–6.
[20] P. M. Narendra and K. Fukunaga, “A branch and bound algorithm for feature subset selection,” *IEEE Transactions on computers*, vol. 9, no. C-26, pp. 917–922, 1977.